Optimality of general lattice transformations with applications to the Bain strain in steel

Article (Accepted Version)

Optimality of General Lattice Transformations with Applications to the Bain Strain in Steel

Konstantinos Koumatos*

Gran Sasso Science Institute,
Viale Fransesco Crispi 7,
67100 L'Aquila, Italy

Anton Muehlemann†

Mathematical Institute, University of Oxford,
Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road,
Oxford OX2 6GG, United Kingdom

Dated: June 24, 2016

This article provides a rigorous proof of a conjecture by E.C. Bain in 1924 on the optimality of the so-called "Bain strain" based on a criterion of least atomic movement. A general framework that explores several such optimality criteria is introduced and employed to show the existence of optimal transformations between any two Bravais lattices. A precise algorithm and a GUI to determine this optimal transformation is provided. Apart from the Bain conjecture concerning the transformation from face-centred cubic to body-centred cubic, applications include the face-centred cubic to body-centred tetragonal transition as well as the transformation between two triclinic phases of Terephthalic Acid.

MSC (2010): 74N05, 74N10

KEYWORDS: lattice transformation, least atomic movement, Bravais lattices, Bain strain in steel, fcc-to-bcc, Terephthalic Acid

*konstantinos.koumatos@gssi.infn.it
†muehlemann@maths.ox.ac.uk

Contents

1 Introduction

In his seminal article on "The nature of martensite" Bain [Bai24] proposed a mechanism that transforms the face-centred cubic (fcc) lattice to the body-centred cubic (bcc) lattice, a phase transformation most importantly manifested in low carbon steels. He writes:

> *"It is reasonable, also, that the atoms themselves will rearrange* [...] *by a method that will require least temporary motion.* [...] *A mode of atomic shift requiring minimum motion was conceived by the author* [...] *"*
>
> Bain, 1924

The key observation that led to his famous correspondence was that *"If one regards the centers of faces as corners of a new unit, a body-centered structure is already at hand; however, it is tetragonal instead of cubic"*. He remarks that this is not surprising *"as it is the only easy method of constructing a bcc atomic structure from the fcc atomic structure"*.

Even though now widely accepted, his mechanism, which he illustrated with a model made of cork balls and needles (see Fig. 1), was not without criticism from his contemporaries. In their fundamental paper Kurdjumov & Sachs [KS30] wrote [free translation from German] that "nothing certain about the mechanism of the martensite transformation is known. Bain imagines that a tetragonal unit cell within the fcc lattice transforms into a bcc unit cell through compression along one direction and expansion along the two other. However a proof of this hypothesis is still missing".[1] Interestingly, without being aware of it, the authors implicitly used the Bain mechanism in their derivation of the Kurdjumov & Sachs orientation relationships (see [KM] for details).

[1] *"Über den Mechanismus dieser ,,Martensitumwandlung" ist bisher nichts Sicheres bekannt. Bain stellt sich vor, daß eine tetra- gonalkörperzentrierte Elementarzelle des Austenits durch Schrumpfung in der einen Richtung und Ausdehnung in den beiden anderen in die kubis-*

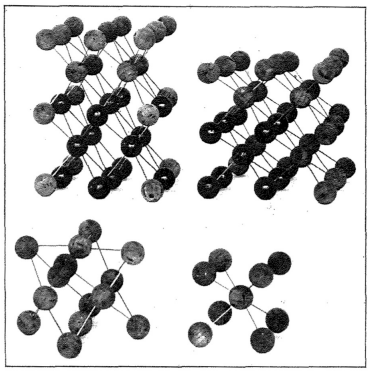

Figure 1: From E.C. Bain: the small models show the fcc and bcc unit cells; the large models represent 35 atoms in an fcc and bcc arrangement respectively.

In subsequent years, the determination of the transformation mechanism remained of great interest. In their paper on "Atomic Displacements in the Austenite-Martensite Transformation" [JW48] Jaswon and Wheeler again acknowledged that

> *"Off all the possible distortions of a primitive unit cell of the face-centred cubic structure, which could generate a body-centred cubic structure of the given relative orientation, the one which actually occurs is the smallest"* Jaswon and Wheeler, 1948

By combining it with experimental observations of the orientation relationships they devised an algorithm to derive the strain tensor. However their approach is only applicable to cases where the orientation relationship is known a-priori.

With the years passing and a number of supporting experimental results (for a discussion see e.g. [BW72]) the Bain mechanism rose from a conjecture to a widely accepted fact. Nevertheless, almost a century after Bain first announced his correspondence a rigorous proof based on the assumption of minimal atom movement has been missing. Of course, the transformation from fcc-to-bcc is not the only instance where the determination of the transformation strain is of interest. The overall question remains the same: Which transformation strain(s), out of all the possible deformations mapping the lattice of the parent phase to the lattice of the product phase, require(s) the least atomic movement?

To provide a definite answer to this question one first needs to quantify the notion of least atomic movement in such a way that it does not require additional input from experiments. Then one needs to establish a framework that singles out the optimal transformation among the infinite number of possible lattice transformation strains. One way to appropriately quantify least atomic movement is the criterion of smallest principal strains as suggested by Lomer in [Lom55]. In his paper, Lomer compared 1600 different lattice correspondences for the β to α phase transition in Uranium and concluded that only one of them involved strains of less than 10%.

chraumzentrierte des α-Eisens übergeht. Eine Bestätigung für diese Anschauung konnte bisher nicht erbracht werden."

More recently, in [CSTJ16] an algorithm is proposed to determine the transformation strain based on a similar minimality criterion (see Remark 6) that also allows for the consideration of different sublattices. The present paper considers a criterion of least atomic movement in terms of a family of different strain measures and, for each such strain measure, rigorously proves the existence of an optimal lattice transformation between *any* two Bravais lattices.[2] As a main application, it is shown that the Bain strain is the optimal lattice transformation from fcc-to-bcc with respect to three of the most commonly used strain measures.

The structure of the paper is as follows: after stating some preliminaries in Section 2 we explore in more depth some mathematical aspects of lattices in Section 3. This section is mainly intended for the mathematically inclined reader and may be skipped on first reading without inhibiting the understanding of Section 4, which constitutes the main part of this paper. In this section we establish a geometric criterion of optimality and prove the existence of optimal lattice transformations for any displacive phase transition between two Bravais lattices. Additionally a precise algorithm to compute these optimal strains is provided. In the remaining subsections, the general theory is applied to prove the optimality of the Bain strain in an fcc-to-bcc transformation, to show that the Bain strain remains optimal in an fcc to body-centred tetragonal (bct) transformation and finally to derive the optimal transformation strain between two triclinic phases of Terephthalic Acid. Similarly to the fcc-to-bcc transition, this phase transformation is of particular interest as it involves large stretches and thus the lattice transformation requiring least atomic movement is not clear.

[2] In particular, no assumptions are made on the type of lattice points (e.g. atoms, molecules) or on the relation between the point groups of the two lattices.

2 Preliminaries

Throughout it is assumed that both the parent and product lattices are Bravais lattices (see Definition 5) and that the transformation strain, i.e. the deformation that maps a unit cell of the parent lattice to a unit cell of the product lattice, is homogeneous.

The following definitions are standard and will be used throughout.

Definition 1. Let $\mathcal{R} \in \{\mathbb{Z}, \mathbb{R}\}$ denote the set of integer or real numbers respectively and define

- $\mathcal{R}^{3\times3}$: Vector space of matrices with entries in $\mathcal{R}$.
- $\mathcal{R}^{3\times3}_{+} := \mathcal{R}^{3\times3} \cap \{\det > 0\}$: Orientation preserving matrices with entries in $\mathcal{R}$.
- $\mathrm{GL}(3,\mathcal{R}) := \mathcal{R}^{3\times3} \cap \{A \in \mathcal{R}^{3\times3} \text{ is invertible}\}$ *(General linear group).*
- $\mathrm{GL}^{+}(3,\mathcal{R}) := \mathrm{GL}(3,\mathcal{R}) \cap \mathcal{R}^{3\times3}_{+}$: Set of orientation preserving invertible 3×3 matrices with entries in $\mathcal{R}$.
- $\mathrm{SL}(3,\mathbb{Z}) := \mathrm{GL}^{+}(3,\mathbb{Z}) = \{A \in \mathbb{Z}^{3\times3} : \det A = 1\}$ *(Special linear group).*
- $\mathrm{SO}(3) = \{R \in \mathbb{R}^{3\times3} : R^{\mathrm{T}}R = \mathbb{I}, \det R = 1\}$ *(Group of proper rotations).*
- $\mathcal{P}^{24} \subset \mathrm{SO}(3)$ *(Symmetry group of a cube)* - see Lemma 1 below.

Further define the multiplication of a matrix F and a set of matrices $\mathcal{S}$ by $F.\mathcal{S} := \{FS : S \in \mathcal{S}\}$.

The following Lemma establishes a characterisation of the group $\mathcal{P}^{24}$, i.e. the set of all rotations that map a cube to itself.

Lemma 1. *Let $Q = [-1,1]^3$ be the cube of side length 2 centred at 0 and define $\mathcal{P}^{24} = \{P \in \mathrm{SO}(3) : PQ = Q\}$. Then $|\mathcal{P}^{24}| = 24$ and $\mathcal{P}^{24} = \mathrm{SO}(3) \cap \mathrm{SL}(3,\mathbb{Z})$.*

Proof. Suppose that $P \in \mathcal{P}^{24}$ and let $\{e_i\}_{i=1,2,3}$ denote the standard basis of $\mathbb{R}^3$. By linearity, P maps the face centres of Q to face centres, i.e. for each $i = 1,2,3$ there exists $j \in \{1,2,3\}$ such that $Pe_i = \pm e_j$. Therefore, $P_{ki} = Pe_i \cdot e_k = \pm\delta_{kj} \in \{-1,0,1\}$ and thus $P \in \mathrm{SL}(3,\mathbb{Z})$.

Conversely, if $P \in \mathrm{SO}(3) \cap \mathrm{SL}(3,\mathbb{Z})$, its columns form an orthonormal basis and since P has integer entries the columns have to be in the set $\{\pm e_i\}_{i=1,2,3}$. Hence, P is of the form

$$P = \sum_{j=1}^{3} \pm e_{k_j} \otimes e_j$$

and $Pe_i = \pm(e_i \cdot e_j)e_{k_j}$. Thus, P maps face centres of Q to face centres and, by linearity, the cube to itself. Further, there are precisely six choices (3×2) for the first column $\pm e_{k_1}$, four choices (2×2) for $\pm e_{k_2}$ and two choices for $\pm e_{k_3}$. Thus, taking into account the determinant constraint, there are 24 elements in $\mathcal{P}^{24}$. □

Remark 1. Essentially the same proof can be used to show that $\mathrm{SO}(n)\cap\mathrm{SL}(n,\mathbb{Z}) = \mathcal{P}^{N_n}$, where $\mathcal{P}^{N_n}$ with $N_n = 2^{n-1}n!$ denotes the symmetry group of a n-dimensional cube. As before the value of N_n arises from having $2\times n$ choices for the first column of Q, then $2\times(n-1)$ choices for the second column of Q, ... and the last column of Q is already fully determined by the determinant constraint.

Definition 2. (Pseudometric and metric)
Let X be a vector space and $x, y, z \in X$. A map $d : X\times X \to [0,\infty)$ is a *pseudometric* if

1. $d(x,x) = 0$,
2. $d(x,y) = d(y,x)$ *(symmetry)*,
3. $d(x,z) \le d(x,y) + d(y,z)$ *(triangle inequality)*.

If in addition d is positive definite, i.e. $d(x,y) = 0 \Leftrightarrow x = y$, we call d a *metric*.

Definition 3. (Matrix norms)
For a given matrix $A \in \mathbb{R}^{3\times3}$ we define the following norms

* *Frobenius norm*:
$$|A| = \sqrt{\mathrm{Tr}(A^{\mathrm{T}}A)} = \sqrt{\sum_{i,j=1}^{3} A_{ij}^2}.$$
* *Spectral norm*:
$$|A|_2 = \sup_{|x|=1} |Ax| = \sqrt{\max_{i=1,2,3} \lambda_i(A^{\mathrm{T}}A)} = \max_{i=1,2,3} \nu_i(A),$$
where for $i = 1,2,3$, $\nu_i(A)$ are the principal stretches/singular values of A and $\lambda_i(A^{\mathrm{T}}A)$ are the eigenvalues of $A^{\mathrm{T}}A$.
* *Column max norm*:
$$\|A\|_{2,\infty} = \max_{i=1,2,3} |Ae_i| = \max_{i=1,2,3} |a_i|,$$
where $\{a_1, a_2, a_3\}$ are the columns of A.

Unless otherwise specified, here and throughout the rest of the paper $|\cdot|$ always denotes the Euclidean norm if the argument is a vector in $\mathbb{R}^3$ and the Frobenius norm if the argument is a matrix in $\mathbb{R}^{3\times3}$. Additionally, we henceforth denote by $\mathrm{col}[A] := \{a_1, a_2, a_3\}$ the columns of the matrix A and then write $A = [a_1, a_2, a_3]$.

The proofs of the following statements are elementary and can be found in standard textbooks on linear algebra (see e.g. [Gen07]).

Lemma 2. *(Properties of matrix norms)*
Both the Frobenius norm and the spectral norm are unitarily equivalent*, that is*

$$|RAS| = |A| \tag{1}$$

for any $R, S \in \mathrm{SO}(3)$. *Further both norms are* sub-multiplicative, *that is given* $A, B, C \in \mathbb{R}^{3\times3}$ *then* $|ABC| \leq |A||B||C|$ *and thus in particular if* $|A||C| \neq 0$ *then*

$$|B| \geq \frac{|ABC|}{|A||C|}. \tag{2}$$

Further the spectral norm is compatible with the Euclidean norm on $\mathbb{R}^3$, *that is*

$$|Ax| \leq |A|_2|x|. \tag{3}$$

The following sets will be of particular importance when proving the optimality of lattice transformations.

Definition 4. ($\mathrm{SL}^{\mathrm{k}}(3,\mathbb{Z})$)
For $k \in \mathbb{N}$ define

$$\mathrm{SL}^{\mathrm{k}}(3,\mathbb{Z}) := \{A \in \mathrm{SL}(3,\mathbb{Z}) : |A_{mn}| \leq k \ \forall\, m,n \in \{1,2,3\}\}$$

and

$$\mathrm{SL}^{-\mathrm{k}}(3,\mathbb{Z}) := \{A \in \mathrm{SL}(3,\mathbb{Z}) : |\left(A^{-1}\right)_{mn}| \leq k \ \forall\, m,n \in \{1,2,3\}\}.$$

Clearly $\mathrm{SL}^{\mathrm{j}}(3,\mathbb{Z}) \subset \mathrm{SL}^{\mathrm{k}}(3,\mathbb{Z}) \ \forall\, 0 \leq j \leq k$ and $|\mathrm{SL}^{-\mathrm{k}}(3,\mathbb{Z})| = |\mathrm{SL}^{\mathrm{k}}(3,\mathbb{Z})|$ for all $k \in \mathbb{Z}$. For example we have $|\mathrm{SL}^1(3,\mathbb{Z})| = 3\,480$, $|\mathrm{SL}^2(3,\mathbb{Z})| = 67\,704$, $|\mathrm{SL}^3(3,\mathbb{Z})| = 640\,824$, $|\mathrm{SL}^4(3,\mathbb{Z})| = 2\,597\,208$, $|\mathrm{SL}^5(3,\mathbb{Z})| = 10\,460\,024$ and $|\mathrm{SL}^6(3,\mathbb{Z})| = 28\,940\,280$.

Below we recall some basic definitions and results from crystallography.

Definition 5. (Bravais lattice, [Bha03] Ch. 3)
Let $F = [f_1, f_2, f_3] \in \mathrm{GL}^+(3,\mathbb{R})$, where $\mathrm{col}[F] = \{f_1, f_2, f_3\}$ are the columns of F. We define the *Bravais lattice* $\mathcal{L}(F)$ *generated by* F as the lattice generated by $\mathrm{col}[F]$, i.e.

$$\mathcal{L}(F) := \mathrm{col}[F.\mathbb{Z}_+^{3\times3}].$$

Thus by definition a Bravais lattice is $\mathrm{span}_{\mathbb{Z}}\{f_1, f_2, f_3\}$ together with an orientation.

Definition 6. (Primitive, base-, body- and face-centred unit cells)
Let $\mathcal{L} = \mathcal{L}(F)$ be generated by $F \in \mathrm{GL}^+(3,\mathbb{R})$. We call the parallelepiped spanned by $\mathrm{col}[F]$ with one atom at each vertex a *primitive unit cell* of the lattice. We call a primitive unit cell with additional atoms in the centre of the bases a *base-centred* unit cell, we call a primitive unit cell with one additional atom in the body centre a *body-centred* (bc) unit cell and we call a primitive unit cell with additional atoms in the centre of each of the faces a *face-centred* (fc) unit cell.

Remark 2. For any lattice generated by a base-, body- or face-centred unit cell there is a primitive unit cell that generates the same lattice. The following table gives the lattice vectors that generate the equivalent primitive unit cell for a given base-centred (C), body-centred (I) or face-centred (F) unit cell spanned by the vectors $\{a, b, c\} \in \mathbb{R}^3$. For our purposes all unit cells that generate the same lattice

primitive (P)	base-centred (C)	body-centred (I)	face-centred (F)
$\{a, b, c\}$	$\{\frac{a-b}{2}, \frac{a+b}{2}, c\}$	$\{\frac{-a+b+c}{2}, \frac{a-b+c}{2}, \frac{a+b-c}{2}\}$	$\{\frac{b+c}{2}, \frac{a+c}{2}, \frac{a+b}{2}\}$

Table 1: Lattice vectors of a primitive unit cell that generates the same lattice.

are equivalent and in order to keep the presentation as simple as possible we will always work with the primitive description of a lattice. However, we note that often in the literature the unit cell is chosen such that it has maximal symmetry.

For example for a face-centred cubic lattice, the unit cell would be chosen as face-centred and spanned by $\mathrm{col}[\mathbb{I}] = \{e_1, e_2, e_3\}$ so that it has the maximal $\mathcal{P}^{24}$ symmetry. However, if one considers primitive unit cells that span the same fcc lattice, the one with maximal symmetry is given by the last entry in Table 1 and thus spanned by $\mathrm{col}[F]$, where

$$F = \frac{1}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

and has only 6-fold symmetry.

Lemma 3. *(Identical lattice bases, [Bha03] Result 3.1)*
Let $\mathcal{L}(F)$ be generated by $F = [f_1, f_2, f_3]$ and $\mathcal{L}(G)$ be generated by $G = [g_1, g_2, g_3]$. Then

$$\mathcal{L}(F) = \mathcal{L}(G) \Leftrightarrow G = F\mu \Leftrightarrow g_i = \sum_{j=1}^{3} \mu_{ji} f_j,$$

for some $\mu \in \mathrm{SL}(3, \mathbb{Z})$. The same result holds for a face-, base- and body-centred unit cell.

Definition 7. (Lattice transformation)
Given two lattices $\mathcal{L}_0 = \mathcal{L}(F)$ and $\mathcal{L}_1 = \mathcal{L}(G)$ generated by $F, G \in \mathrm{GL}^+(3, \mathbb{R})$ we call any matrix $H \in \mathrm{GL}^+(3, \mathbb{R})$ such that $H.\mathcal{L}_0 = \mathcal{L}_1$ a *lattice transformation* from $\mathcal{L}_0$ to $\mathcal{L}_1$.

Remark 3. In the terminology of Ericksen (see e.g. [PZ02] p. 62ff.), if $\mathcal{L}_0 = \mathcal{L}_1$ (i.e. $G = F\mu$), the matrices H in Definition 7 are precisely all the orientation preserving elements in the *global symmetry group* of F. Additionally, the matrices H that are also rotations constitute the *point group* of the lattice. We point out that, in this terminology, $\mathcal{P}^{24}$ is the point group of any cubic lattice.

We end this section by defining the atom density of the lattice $\mathcal{L}(F)$ and relating it to the determinant of F.

Definition 8. (Atom density)
For a given lattice $\mathcal{L}$ we define the atom density $\rho(\mathcal{L})$ by

$$\rho(\mathcal{L}) := \lim_{N\to\infty} \frac{\#\{Q_N \cap \mathcal{L}\}}{N^3},$$

where $Q_N = [0, N]^3$ is the cube with side-length N and $\#$ counts the number of elements in a discrete set. Thus $\rho(\mathcal{L})$ is the average number of atoms per unit volume.

Lemma 4. *Let $\mathcal{L} = \mathcal{L}(F)$ be generated by $F \in \mathrm{GL}^+(3, \mathbb{R})$ then*

$$\rho(\mathcal{L}) = \frac{1}{\det F}.$$

In particular a transformation $H = GF^{-1}$ from $\mathcal{L}_0 = \mathcal{L}(F)$ to $\mathcal{L}_1 = \mathcal{L}(G)$ does not change the atom density if and only if it is volume preserving, i.e. $\det H = 1$. Further the atom density is well defined.

Proof. Denote by $\mathcal{U} \subset \mathbb{R}^3$ the unit cell spanned by $\mathrm{col}[F]$, so that the volume of $\mathcal{U}$ is given by $|\mathcal{U}| = \det F$. Taking n distinct points $x_i \in \mathcal{L}$ we find that $|\bigcup_{i=1}^n (x_i + \mathcal{U})| = n \det F$ since all elements are disjoint (up to zero measure). Let l denote the side length of the smallest cube that contains $\mathcal{U}$. Further, as in Definition 8, let $Q_N = [0, N]^3$ denote a cube of side-length N and define $Q_N^{\pm} = [\mp 2l, N \pm 2l]^3$. Then

$$Q_N^- \subset \bigcup_{x \in \mathcal{L} \cap Q_N} (x + \mathcal{U}) \subset Q_N^+$$

and thus by taking the volumes of the sets we obtain

$$(N - 4l)^3 \leq \#\{Q_N \cap \mathcal{L}\} \det F \leq (N + 4l)^3.$$

Dividing by N^3 and taking the limit $N \to \infty$ yields the result. Since this limit exists for all sequences $N \to \infty$ the density is well defined. □

3 Metrics and equivalence on matrices and lattices

Definition 9. (Equivalent matrices and lattices)
We define an equivalence relation $\sim$ between matrices F, G in $\mathrm{GL}^+(3, \mathbb{R})$ by

$$F \sim G :\Leftrightarrow \exists R \in \mathrm{SO}(3) : G = RF,$$

so that the equivalence class $[F]$ of F is given by $[F] = \{G \in \mathbb{R}^{3\times 3}_+ : F \sim G\}$. We denote the quotient space, i.e. the space of all equivalence classes, by

$$\overline{\mathrm{GL}^+(3, \mathbb{R})} := \{[F] : F \in \mathrm{GL}^+(3, \mathbb{R})\}.$$

Furthermore, we define an equivalence relation $\sim$ between two lattices $\mathcal{L}_0$ and $\mathcal{L}_1$ by

$$\mathcal{L}_0 \sim \mathcal{L}_1 :\Leftrightarrow \exists R \in \mathrm{SO}(3) : \mathcal{L}_1 = R.\mathcal{L}_0.$$

We are now in a position to define a metric on the quotient spaces.

Lemma and Definition 10. *(Induced metric)*
Any pseudometric $d : \mathrm{GL}^+(3,\mathbb{R}) \times \mathrm{GL}^+(3,\mathbb{R}) \to [0,\infty)$ *with the property*

$$d(F,G) = 0 \Leftrightarrow G = T^* F \text{ for some } T^* \in \mathrm{SO}(3) \qquad (*)$$

naturally induces a metric $\bar{d}$ *on* $\overline{\mathrm{GL}^+(3,\mathbb{R})}$ *via*

$$\bar{d}([F],[G]) = \min_{R,S \in \mathrm{SO}(3)} d(RF,SG).$$

Proof. The quantity $\bar{d}$ is clearly well defined, so that in particular $\bar{d}([F],[G]) = \bar{d}(F,G)$ and we may henceforth drop the $[\cdot]$ in the arguments of $\bar{d}$. We first show that $\bar{d}$ is a metric on $\overline{\mathrm{GL}^+(3,\mathbb{R})}$. Positivity and symmetry are obvious from the definition. For definiteness note that if $\min_{R,S\in\mathrm{SO}(3)} d(RF,SG) = 0$ then by $(*)$ we have $S^*G = T^*R^*F$ for some $R^*, S^*, T^* \in \mathrm{SO}(3)$ and thus $F \sim G$. It remains to show the triangle inequality. We have

$$\begin{aligned}
\bar{d}(F,H) &= \min_{R,S\in\mathrm{SO}(3)} d(RF,SH) \\
&\le \min_{R,S} \left(d(RF,G) + d(G,SH)\right) \\
&\le \min_{R,S'} \left(d(RF,S'G) + \cancel{d(G,S'G)}\right) + \min_{R',S} \left(d(SH,R'G) + \cancel{d(G,R'G)}\right) \\
&= \bar{d}(F,G) + \bar{d}(G,H),
\end{aligned}$$

where we have used the triangle inequality, symmetry of d and $(*)$. □

Example 1. The family of maps $d_r : \mathrm{GL}^+(3,\mathbb{R}) \times \mathrm{GL}^+(3,\mathbb{R}) \to [0,\infty)$, $r \in \mathbb{R}\backslash\{0\}$ given by

$$d_r(F,G) = |(F^{\mathrm{T}}F)^{r/2} - (G^{\mathrm{T}}G)^{r/2}|$$

are pseudometrics such that $(*)$ holds. In particular, each of them induces a metric $\bar{d}_r$ on the quotient space $\overline{\mathrm{GL}^+(3,\mathbb{R})}$.

Proof. Positivity is obvious. The triangle inequality follows from the corresponding property of the Frobenius norm, i.e.

$$\begin{aligned}
d_r(F,H) &= |(F^{\mathrm{T}}F)^{r/2} - (H^{\mathrm{T}}H)^{r/2}| \\
&\le |(F^{\mathrm{T}}F)^{r/2} - (G^{\mathrm{T}}G)^{r/2}| + |(G^{\mathrm{T}}G)^{r/2} - (H^{\mathrm{T}}H)^{r/2}| = d_r(F,G) + d_r(G,H)
\end{aligned}$$

and the property $(*)$ follows from

$$\begin{aligned}
d_r(F,G) = 0 &\Leftrightarrow (F^{\mathrm{T}}F)^{r/2} = (G^{\mathrm{T}}G)^{r/2} \Leftrightarrow (FG^{-1})^{\mathrm{T}}(FG^{-1}) = \mathbb{I} \\
&\Leftrightarrow FG^{-1} \in \mathrm{SO}(3) \Leftrightarrow F = T^*G \text{ for some } T^* \in \mathrm{SO}(3).
\end{aligned}$$

□

Remark 4. The metric d_2 has previously been used in [Bha03, Chapter 3], where it was defined as the distance between the metric $C_F = F^{\mathrm{T}}F = (f_i \cdot f_j)_{ij}$ of a set of lattice vectors $\mathrm{col}[F] = \{f_1, f_2, f_3\}$ and the metric $C_G = G^{\mathrm{T}}G = (g_i \cdot g_j)_{ij}$ of a set of lattice vectors $\mathrm{col}[G] = \{g_1, g_2, g_3\}$. The use of the term metric in [Bha03] is not to be confused with the use of metric in the present paper.

4 Optimal lattice transformations

This section embodies the main part of the present paper. We first establish what we mean by an optimal transformation from one lattice to another and then, for a family of such criteria, show the existence of optimal transformations between *any* two Bravais lattices.

Lemma 5. *Let $\mathcal{L}_0 = \mathcal{L}(F)$ and $\mathcal{L}_1 = \mathcal{L}(G)$ be generated by $F, G \in \mathrm{GL}^+(3,\mathbb{R})$. Then all possible lattice transformations from $\mathcal{L}_0$ to $\mathcal{L}_1$ are given by $H_\mu = G\mu F^{-1}, \mu \in \mathrm{SL}(3,\mathbb{Z})$. In particular, the lattices coincide if and only if there exists $\mu \in \mathrm{SL}(3,\mathbb{Z})$ such that $G\mu F^{-1} = \mathbb{I}$ and they are equivalent, i.e. $\mathcal{L}_0 \sim \mathcal{L}_1$, if and only if there exists $\mu \in \mathrm{SL}(3,\mathbb{Z})$ such that $G\mu F^{-1} \in \mathrm{SO}(3)$.*

Proof. Let $H_\mu = G\mu F^{-1}, \mu \in \mathrm{SL}(3,\mathbb{Z})$. Then

$$\begin{aligned} H_\mu.\mathcal{L}_0 = H_\mu.\mathcal{L}(F) = H_\mu.\mathrm{col}[F.\mathbb{Z}_+^{3\times3}] = \mathrm{col}[H_\mu F.\mathbb{Z}_+^{3\times3}] \\ = \mathrm{col}[G\mu F^{-1}F.\mathbb{Z}_+^{3\times3}] = \mathrm{col}[G\mu.\mathbb{Z}_+^{3\times3}] = \mathrm{col}[G.\mathbb{Z}_+^{3\times3}] = \mathcal{L}_1, \end{aligned}$$

where we used that μ is invertible over $\mathbb{Z}$, so that $\mu.\mathbb{Z}_+^{3\times3} = \mathbb{Z}_+^{3\times3}$. Thus H_μ is a lattice transformation from $\mathcal{L}_0$ to $\mathcal{L}_1$. For the reverse direction we know by Lemma 3 that $\mathcal{L}(F) = \mathcal{L}(F')$ if and only if $F' = F\mu', \mu' \in \mathrm{SL}(3,\mathbb{Z})$ and $\mathcal{L}(G) = \mathcal{L}(G')$ if and only if $G' = G\mu'', \mu'' \in \mathrm{SL}(3,\mathbb{Z})$ so that all possible generators for $\mathcal{L}_0$ are given by $F\mu', \mu' \in \mathrm{SL}(3,\mathbb{Z})$ and all possible generators for $\mathcal{L}_1$ are given by $G\mu'', \mu'' \in \mathrm{SL}(3,\mathbb{Z})$. Thus any lattice transformation from $\mathcal{L}_0$ to $\mathcal{L}_1$ is given by

$$H_{\mu'}^{\mu''} = G\mu''\mu'^{-1}F^{-1}, \quad \mu', \mu'' \in \mathrm{SL}(3,\mathbb{Z}).$$

But by the group property we may set $H_\mu := H_{\mu'}^{\mu''}$ with $\mu = \mu''\mu'^{-1} \in \mathrm{SL}(3,\mathbb{Z})$. □

Definition 11. (d-optimal lattice transformations)
Given two lattices $\mathcal{L}_0 = \mathcal{L}(F)$ and $\mathcal{L}_1 = \mathcal{L}(G)$ with $F, G \in \mathrm{GL}^+(3,\mathbb{R})$ and a metric d, we call a lattice transformation $H_\mu = G\mu F^{-1}$, $\mu \in \mathrm{SL}(3,\mathbb{Z})$ (cf. Lemma 5) *d-optimal* if it minimises the distance to $\mathbb{I}$ with respect to d, i.e. $H_{\min} = G\mu_{\min}F^{-1}$, where

$$\mu_{\min} = \underset{\mu\in\mathrm{SL}(3,\mathbb{Z})}{\arg\min}\, d(H_\mu, \mathbb{I}). \tag{4}$$

If d is a pseudometric satisfying $(*)$, we call a transformation *d-optimal* if it is optimal with respect to the induced metric $\bar{d}$ on the quotient space $\overline{\mathrm{GL}^+(3,\mathbb{R})}$, i.e.

the d-optimal transformation is the one that is d-closest to being a pure rotation and maps $\mathcal{L}_0$ to a lattice in the equivalence class $[\mathcal{L}_1] \coloneqq \{\mathcal{L} : \mathcal{L} \sim \mathcal{L}_1\}$ of $\mathcal{L}_1$.

Example 2. For the pseudometrics d_r from Example 1 the explicit expressions for the distance in (4) read

$$d_r(H, \mathbb{I}) = |(H^{\mathrm{T}}H)^{r/2} - \mathbb{I}| = \left(\sum_{i=1}^{3} (\nu_i^r - 1)^2 \right)^{1/2}, \tag{5}$$

where ν_i, $i = 1, 2, 3$ are the principal stretches/singular values of H. The quantities $(H^{\mathrm{T}}H)^{r/2} - \mathbb{I}$ are clearly measures of principal strain and are known as the Doyle-Ericksen strain tensors (see [DE56, p. 65]). For $r = 1$, it is simple to verify that

$$d_1(H, \mathbb{I}) = \operatorname{dist}(H, \mathrm{SO}(3)) = \min_{R \in \mathrm{SO}(3)} |H - R|.$$

Remark 5. For a general metric d as in Definition 11, the optimal transformation $H_{\min}$ between $\mathcal{L}(F)$ and $\mathcal{L}(G)$ is unchanged under actions of the point groups of both lattices, i.e.

$$H_{\min} = G \operatorname*{arg\,min}_{\mu \in \mathrm{SL}(3,\mathbb{Z})} d(G\mu F^{-1}, \mathbb{I}) F^{-1} = (PG) \operatorname*{arg\,min}_{\mu \in \mathrm{SL}(3,\mathbb{Z})} d((PG)\mu(QF)^{-1}, \mathbb{I})(QF)^{-1}$$

for any P in the point group of $\mathcal{L}(G)$ and Q in the point group of $\mathcal{L}(F)$. Throughout the rest of the paper, we only use pseudometrics satisfying $(*)$. In this case, the notion of optimality is not only invariant under actions of the respective point groups but also under rigid body rotations of the product lattice. Thus Definition 11 returns an equivalence class $[H_{\min}] \in \overline{\mathrm{GL}^+(3, \mathbb{R})}$ of d-optimal transformations. By the polar decomposition theorem we may, and henceforth always will, pick the symmetric representative

$$\bar{H}_{\min} \coloneqq \sqrt{H_{\min}^{\mathrm{T}} H_{\min}} \in \mathbb{R}^{3\times 3}_{\mathrm{sym}},$$

i.e. the pure *stretch component* of the transformation $H_{\min}$. Note that in general the set of minimising equivalence classes $\{[H_{\min}] : H_{\min} \text{ is } d\text{-optimal}\}$ may contain more than one element. In such a case, different regions of the parent lattice may transform according to any of these optimal strains, giving rise to e.g. twinning.

The pseudometrics d_r are additionally invariant under rotations from the right, i.e. $d_r(H, \mathbb{I}) = d_r(HS, \mathbb{I})$ for all $S \in \mathrm{SO}(3)$. For any such pseudometric, a rigid body rotation R of the parent lattice $\mathcal{L}_0$ results in an optimal transformation with stretch component $R\bar{H}_{\min}R^{\mathrm{T}}$, where $\bar{H}_{\min}$ is the stretch component of the optimal transformation from $\mathcal{L}_0$ to $[\mathcal{L}_1]$. Note that $R\bar{H}_{\min}R^{\mathrm{T}}$ is simply $\bar{H}_{\min}$ expressed in a different basis and in particular, even though the coordinate representation is different, the underlying transformation mechanism is unchanged.

Our main theorem says that a d_r-optimal lattice transformation always exists and the following lemma will be the crucial tool in the proof.

Key Lemma 6. *Let H be a lattice transformation from $\mathcal{L}_0 = \mathcal{L}(F)$ to $\mathcal{L}_1 = \mathcal{L}(G)$ and consider a lattice vector $f \in \mathcal{L}_0$ that is transformed by H to $g = Hf$. Then*

$$\nu_{\max}(H) \geq |g|/|f| \geq \nu_{\min}(H),$$

where $\nu_{\min}(H)$, $\nu_{\max}(H)$ denote, respectively, the smallest and largest principal stretches/singular values of H. In particular for any $s > 0$,

$$\mathrm{d}_s(H, \mathbb{I}) = \left(\sum_{i=1}^{3} (\nu_i^s - 1)^2 \right)^{1/2} \geq \max_i \frac{|Hf_i|^s}{|f_i|^s} - 1, \tag{6}$$

$$\mathrm{d}_{-s}(H, \mathbb{I}) = \left(\sum_{i=1}^{3} (\nu_i^{-s} - 1)^2 \right)^{1/2} \geq \max_i \frac{|H^{-1}g_i|^s}{|g_i|^s} - 1, \tag{7}$$

where $\mathrm{col}[F] = \{f_1, f_2, f_3\}$ and $\mathrm{col}[G] = \{g_1, g_2, g_3\}$.

Proof. Consider the singular value decomposition of $H = UDV$, where $D = \mathrm{diag}(\nu_1, \nu_2, \nu_3)$ and $U, V \in \mathrm{SO}(3)$. Then

$$|g| = |UDVf| = |DVf| \leq \max_i \nu_i(H) |Vf| = \nu_{\max}(H)|f|$$

and analogously for the lower bound. □

Theorem 1. *Given two lattices $\mathcal{L}_0 = \mathcal{L}(F)$ and $\mathcal{L}_1 = \mathcal{L}(G)$ generated by $F, G \in \mathrm{GL}^+(3, \mathbb{R})$ respectively, there exists a d_r-optimal lattice transformation $H_{\mu_{\min}} = G\mu_{\min}F^{-1}$, for any $r \in \mathbb{R}\backslash\{0\}$. For $s > 0$ all optimal changes of basis are contained in the finite compact sets*

$$d_s: \quad \left\{ \mu \in \mathrm{SL}(3, \mathbb{Z}) : \|\mu\|_{2,\infty}^s \leq \frac{\|F\|_{2,\infty}^s}{\nu_{\min}^s(G)} (m_{0,s} + 1) \right\}, \tag{8}$$

$$d_{-s}: \quad \left\{ \mu \in \mathrm{SL}(3, \mathbb{Z}) : \|\mu^{-1}\|_{2,\infty}^s \leq \frac{\|G\|_{2,\infty}^s}{\nu_{\min}^s(F)} (m_{0,-s} + 1) \right\}, \tag{9}$$

where $\nu_{\min}(A)$ denotes the smallest principal stretch/singular value of A and

$$m_{0,r} \coloneqq d_r(H_{\mathbb{I}}, \mathbb{I}) = d_r(GF^{-1}, \mathbb{I}).$$

Proof. As the minimisation is over the discrete set $\mathrm{SL}(3, \mathbb{Z})$ it suffices to show that the minimum is attained in a (compact) finite subset of $\mathrm{SL}(3, \mathbb{Z})$ given by (8) and (9). Let $H_\mu = G\mu F^{-1}$, $\mu \in \mathrm{SL}(3, \mathbb{Z})$ be a lattice transformation from $\mathcal{L}_0 = \mathcal{L}(F)$ to $\mathcal{L}_1 = \mathcal{L}(G)$. Then, letting $\{e_i\}_{i=1,2,3}$ denote the standard basis vectors of $\mathbb{R}^{3\times 3}$,

$$H_\mu f_i = G\mu F^{-1} F e_i = G\mu e_i \text{ and } H_\mu^{-1} g_i = F\mu^{-1} G^{-1} G e_i = F\mu^{-1} e_i$$

and thus, by using the Key Lemma 6 and (3), we obtain

$$\mathrm{d}_s(H_\mu,\mathbb{I}) \geq \max_i \frac{|G\mu e_i|^s}{|f_i|^s} - 1 \quad \geq \frac{\|\mu\|_{2,\infty}^s}{|G^{-1}|_2^s \|F\|_{2,\infty}^s} - 1 = \frac{\nu_{\min}^s(G)\|\mu\|_{2,\infty}^s}{\|F\|_{2,\infty}^s} - 1,$$
$$\mathrm{d}_{-s}(H_\mu,\mathbb{I}) \geq \max_i \frac{|F\mu^{-1}e_i|^s}{|g_i|^s} - 1 \geq \frac{\|\mu^{-1}\|_{2,\infty}^s}{|F^{-1}|_2^s \|G\|_{2,\infty}^s} - 1 = \frac{\nu_{\min}^s(F)\|\mu^{-1}\|_{2,\infty}^s}{\|G\|_{2,\infty}^s} - 1,$$

where in the equality we have used that $\{\nu_i(A^{-1})\}_{i=1,2,3} = \{(\nu_i(A))^{-1}\}_{i=1,2,3}$. Thus $d_r(H_\mu,\mathbb{I}) > d_r(H_\mathbb{I},\mathbb{I})$ for all μ in the complement of the respective sets given by (8) and (9) and therefore H_μ cannot be d_r-optimal. □

Remark 6. The distance $d_1(H,\mathbb{I})$ seems to be the most natural candidate to determine the transformation requiring least atomic movement. The quantities $\nu_i - 1$ measure precisely the displacement along the principal axes and thus their use is in line with the criterion of smallest principal strains as in e.g. [Bha01], [BW72] and [Lom55]. The distance $d_2(H,\mathbb{I})$ seems natural from a mathematical perspective as the tensor $H^{\mathrm{T}}H$ corresponds to the flat metric induced by the deformation H and it has also been used to define the Ericksen-Pitteri neighbourhood of a lattice (see e.g. (2.17) in [BJ92]). Finally, the distance $d_{-2}(H,\mathbb{I})$ has recently been used in [CSTJ16] in order to avoid singular behaviour when considering sublattices.

Below we illustrate the differences of d_1, d_2 and d_{-2} through a simple but instructive 1D example.

Example 3. (A comparison of different optimality conditions)
We consider two atoms A, B that are originally at unit distance, i.e. $|A-B| = 1$ and then move the atom B to its deformed position B'. Thus H is simply a scalar quantity given by $H = |A-B'|/|A-B| = |A-B'|$. It can be seen from Table 2 that

r	$B'-B = 0.5$, $H = \lvert A-B'\rvert = 1.5$	$B'-B = -0.5$, $H = \lvert A-B'\rvert = 0.5$	deformation y such that $d_r(y,\mathbb{I}) = d_r(x,\mathbb{I})$
1	$d_1(H,\mathbb{I}) = 0.5$	$d_1(H,\mathbb{I}) = 0.5$	$y = 2 - x$
2	$d_2(H,\mathbb{I}) = 1.25$	$d_2(H,\mathbb{I}) = 0.75$	$y = \sqrt{2-x^2}$
-2	$d_{-2}(H,\mathbb{I}) = 0.\bar{5}$	$d_{-2}(H,\mathbb{I}) = 3$	$y = \frac{1}{\sqrt{2-x^{-2}}}$

Table 2: Comparison of different distances

d_1 depends only on the distance between B and B'; an expansion by 100% has the same d_1 distance to $\mathbb{I}$ as a contraction to 0, i.e. moving A onto B. The metric d_2 penalises expansions more than contractions; e.g. an expansion by $\approx$ 141% has the same d_2 distance to 1 as a contraction to 0. The metric d_{-2} penalises contractions significantly more than expansions; e.g. an expansion by ∞ has the same d_{-2} distance to $\mathbb{I}$ as a contraction to $\approx 70\%$, i.e. reducing the distance between A and B by $\approx 30\%$.

A remark on the computation of the optimal transformation

Theorem 1 provides the necessary compactness result to reduce the original minimisation problem over the infinite set $\mathrm{SL}(3,\mathbb{Z})$ to a finite subset given by (8) and (9) respectively. To this end, it is worth noting that the smaller the deformation distance $m_{0,r} = d_r(GF^{-1}, \mathbb{I})$ of the initial lattice basis the smaller the radius of the ball in $\mathrm{SL}(3,\mathbb{Z})$ that contains the optimal μ. However, in specific cases, where better estimates are available, it might be advantageous to start with an initial lattice basis that is not optimal.

Nevertheless, in order to explicitly determine the optimal transformations one still needs to compare the distances $d_r(H_\mu, \mathbb{I})$ for all elements contained in the finite sets given by (8) and (9) respectively. This can easily be carried out with any modern computer algebra program and possible implementations can be found in the Appendix.

In order to ensure that the solution of this finite minimisation problem is correct one needs to verify that the difference Δ between the minimal and the second to minimal deformation distance is large compared to possible rounding errors (if any). The computations in Sections 4.1 and 4.2 for the Bain strain from fcc-to-bcc/bct are exact and thus without rounding errors. In Section 4.3 regarding the optimal transformation in Terephthalic Acid we find that $\Delta > 0.015$ which is large compared to machine precision.

4.1 The Bain strain in fcc-to-bcc

Having established the general theory of optimal lattice transformations we apply these results to prove the optimality of the Bain strain with respect to the three different lattice metrics d_r, $r = -2, 1, 2$, from the previous example. In these cases we rigorously prove the optimality of the Bain strain first proposed in [Bai24].

Theorem 2. *(Bain Optimality)*
In a transformation from an fcc to a bcc lattice with no change in atom density, there are three distinct equivalence classes of d_r-optimal lattice transformations for $r = 1, 2, -2$. The stretch components are given by

$$\bar{H}_{\min} \in \{\mathrm{diag}(2^{-1/3}, 2^{1/6}, 2^{1/6}),\ \mathrm{diag}(2^{1/6}, 2^{-1/3}, 2^{1/6}),\ \mathrm{diag}(2^{1/6}, 2^{1/6}, 2^{-1/3})\},$$

i.e. the three Bain strains are the d_r-optimal lattice transformations in a volume preserving fcc-to-bcc transformation for $r = 1, 2, -2$. The respective minimal metric distances are

$$m_{\min,1} = d_1(H_{\min}, \mathbb{I}) = \sqrt{\left(2^{-1/3} - 1\right)^2 + 2\left(2^{1/6} - 1\right)^2} \approx 0.269,$$

$$m_{\min,2} = d_2(H_{\min}, \mathbb{I}) = \sqrt{\left(2^{-2/3} - 1\right)^2 + 2\left(2^{1/3} - 1\right)^2} \approx 0.522,$$

$$m_{\min,-2} = d_{-2}(H_{\min}, \mathbb{I}) = \sqrt{\left(2^{2/3} - 1\right)^2 + 2\left(2^{-1/3} - 1\right)^2} \approx 0.656.$$

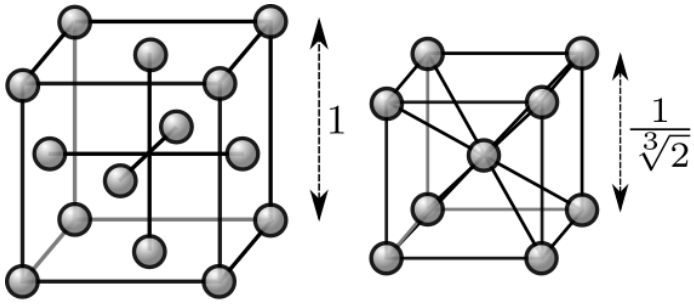

Figure 2: Face-centred and body-centred cubic unit cells with equal atom density.

Proof. Let $\mathcal{L}_0$ denote the fcc lattice, where the fcc unit cell has unit volume and let $\mathcal{L}_1$ denote the bcc lattice with the same atom density (see Fig. 2). Then $\mathcal{L}_0 = \mathcal{L}(F)$ and $\mathcal{L}_1 = \mathcal{L}(B)$, where

$$F = \frac{1}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = [f_1, f_2, f_3], \quad B = 2^{-1/3}\frac{1}{2}\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} = [b_1, b_2, b_3] \tag{10}$$

and, in particular, $\det F = \det B = 4^{-1}$. Let $H_\mu = B\mu F^{-1}$, $\mu \in \mathrm{SL}(3,\mathbb{Z})$ denote the lattice transformation from $\mathcal{L}_0$ to $\mathcal{L}_1$ (cf. Lemma 5). By definition H_μ is optimal if μ satisfies (4). We first show the optimality with respect to d_1 and d_2. We may find an upper bound on the minimum by only considering $\mu \in \mathrm{SL}^1(3,\mathbb{Z})$ (cf. Definition 4). With the help of a computer we find exactly 72 such μ's and all corresponding deformations are (volume preserving) Bain strains. To complete the proof we employ our key Lemma 6 to show that any $\mu \in \mathrm{SL}(3,\mathbb{Z})\backslash\mathrm{SL}^1(3,\mathbb{Z})$ cannot be optimal with respect to either d_1 or d_2. Let $\mu \in \mathbb{Z}^{3\times 3}$ be given by

$$\mu = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix}. \tag{11}$$

Then $b_{\mu,i} = H_\mu f_i = B\mu e_i = \alpha_i b_1 + \beta_i b_2 + \gamma_i b_3$ and after dropping the index i we obtain

$$|b_\mu|^2 = (\alpha^2 + \beta^2 + \gamma^2)|b_1|^2 + 2(\alpha\beta + \beta\gamma + \alpha\gamma)\langle b_1, b_2\rangle, \tag{12}$$

where we have used that $|b_i| = |b_j|$ and $\langle b_i, b_j\rangle = \langle b_k, b_l\rangle$ for all $i \neq j$ and $k \neq l$. We compute $|f_i| = 2^{-1/2}$, $|b_i|^2 = 3\cdot 2^{-8/3}$ and $\langle b_i, b_j\rangle = -2^{-8/3}$. By (6) we estimate

$$d_1(H_\mu, \mathbb{I}) \geq \frac{2^{-4/3}\left(\rho(\alpha,\beta,\gamma)\right)^{1/2}}{2^{-1/2}} - 1, \tag{13}$$

$$d_2(H_\mu, \mathbb{I}) \geq \frac{2^{-8/3}\rho(\alpha,\beta,\gamma)}{2^{-1}} - 1 \tag{14}$$

where $\rho(\alpha,\beta,\gamma) := \alpha^2+\beta^2+\gamma^2+(\alpha-\beta)^2+(\beta-\gamma)^2+(\alpha-\gamma)^2$. If $\mu \in \mathrm{SL}(3,\mathbb{Z})\backslash\mathrm{SL}^1(3,\mathbb{Z})$ then $\rho(\alpha,\beta,\gamma) \geq 8$ and thus

$$d_1(H_\mu,\mathbb{I}) \geq 2^{2/3}-1 > 0.5 \gg m_{\min,1} \text{ and } d_2(H_\mu,\mathbb{I}) \geq 2^{4/3}-1 > 1.5 \gg m_{\min,2}.$$

To show d_{-2}-optimality we consider $H_\mu = B(F\mu^{-1})^{-1}$ and use the ansatz (11) for μ^{-1} instead of μ. We compute

$$|F\mu^{-1}e_i|^2 = \frac{1}{4}\left((\alpha+\beta)^2+(\beta+\gamma)^2+(\alpha+\gamma)^2\right)$$

and we note that $b_i = H_\mu F\mu^{-1}e_i$. Thus by (7) we can estimate

$$d_{-2}(H_\mu,\mathbb{I}) \geq \frac{\frac{1}{4}\sigma(\alpha,\beta,\gamma)}{3\cdot 2^{-8/3}} - 1 = 2\cdot 2^{2/3} - 1 > 2.17 \gg m_{\min,-2}, \tag{15}$$

where $\sigma(\alpha,\beta,\gamma) := (\alpha+\beta)^2+(\beta+\gamma)^2+(\alpha+\gamma)^2$ and we have used that $\sigma(\alpha,\beta,\gamma) \geq 6$ for $\mu^{-1} \in \mathrm{SL}(3,\mathbb{Z})\backslash\mathrm{SL}^1(3,\mathbb{Z})$. Therefore the d_{-2}-optimal μ is contained in $\mathrm{SL}^{-1}(3,\mathbb{Z})$. □

Corollary 1. *The three Bain strains remain the d_r-optimal lattice transformations, $r = 1,2,-2$, from fcc-to-bcc if the volume changes by λ^3, provided that $\lambda > 0.84$ for $r = 1$, $\lambda > 0.64$ for $r = 2$ and $\lambda < 1.19$ for $r = -2$. The stretch components of the three optimal equivalence classes are given by*

$$\bar{H}^\lambda_{\min} \in \{\lambda\,\mathrm{diag}(2^{-1/3},2^{1/6},2^{1/6}), \lambda\,\mathrm{diag}(2^{1/6},2^{-1/3},2^{1/6}), \lambda\,\mathrm{diag}(2^{1/6},2^{1/6},2^{-1/3})\}.$$

The minimal metric distances are given by

$$\begin{aligned}
m^\lambda_{\min,1} &= d_1(H^\lambda_{\min},\mathbb{I}) = \sqrt{\left(2^{-1/3}\lambda-1\right)^2+2\left(2^{1/6}\lambda-1\right)^2},\\
m^\lambda_{\min,2} &= d_2(H^\lambda_{\min},\mathbb{I}) = \sqrt{\left(2^{-2/3}\lambda^2-1\right)^2+2\left(2^{1/3}\lambda^2-1\right)^2},\\
m^\lambda_{\min,-2} &= d_{-2}(H^\lambda_{\min},\mathbb{I}) = \sqrt{\left(2^{2/3}\lambda^{-2}-1\right)^2+2\left(2^{-1/3}\lambda^{-2}-1\right)^2}.
\end{aligned}$$

Proof. Replace $\mu \to \lambda\mu$ in (11) in the proof of Theorem 2. Then (13), (14) and (15) respectively read

$$\begin{aligned}
d_1(H^\lambda_\mu,\mathbb{I}) &\geq \frac{2^{-4/3}\left(\lambda^2\rho(\alpha,\beta,\gamma)\right)^{1/2}}{2^{-1/2}} - 1 \geq \frac{2^{1/6}}{2}\lambda\cdot\inf_{\mathcal{S}}\rho^{1/2} - 1,\\
d_2(H^\lambda_\mu,\mathbb{I}) &\geq \frac{2^{-8/3}\lambda^2\rho(\alpha,\beta,\gamma)}{2^{-1}} - 1 \geq 2^{-2/3}\lambda^2\cdot\inf_{\mathcal{S}}\rho - 1,\\
d_{-2}(H^\lambda_\mu,\mathbb{I}) &\geq \frac{\frac{1}{4}\sigma(\alpha,\beta,\gamma)}{3\cdot 2^{-8/3}\lambda^2} - 1 \geq \frac{2^{2/3}}{3\lambda^2}\cdot\inf_{\mathcal{S}}\sigma - 1.
\end{aligned}$$

If as above $\mathcal{S} = \mathrm{SL}(3,\mathbb{Z})\backslash\mathrm{SL}^1(3,\mathbb{Z})$ then $\inf_{\mathcal{S}}\rho = 8$ and thus $d_1(H^\lambda_\mu,\mathbb{I}) \geq m^\lambda_{\min,1}$ for $\lambda > 0.84$ and $d_2(H^\lambda_\mu,\mathbb{I}) \geq m^\lambda_{\min,2}$ for $\lambda > 0.64$ and $\inf_{\mathcal{S}}\sigma = 6$ so that $d_{-2}(H^\lambda_\mu,\mathbb{I}) \geq m^\lambda_{\min,-2}$ for $\lambda < 1.19$. □

The following Remark concerns the relationship between the 72 minimising states.

Remark 7. (Relations between the minimal deformations for fcc-to-bcc)
Let $\mathcal{L}(F)$ and $\mathcal{L}(B)$ be the fcc and bcc lattices respectively. Let μ_0 be one of the optimal changes of lattice basis and $H_i = B\mu_i F^{-1}, \quad i = 0, \dots, 71$ denote the 72 optimal lattice deformations associated to the optimal changes of basis $\mu_i \in \mathrm{SL}(3,\mathbb{Z})$ given by Theorem 2. Then all optimal H_i's and corresponding μ_i's are given by

$$H_{PQ} = PH_0Q = B\mu_{PQ}F^{-1}, \quad P, Q \in \mathcal{P}^{24},$$

where $\mu_{PQ} = B^{-1}PB\mu_0F^{-1}QF \in \mathrm{SL}(3,\mathbb{Z})$. We note that the latter equation holds since $\mathcal{P}^{24}$ is the point group of both cubic lattices and thus $B^{-1}PB$ and $F^{-1}QF$ are contained in $\mathrm{SL}(3,\mathbb{Z})$. Since there are only three equivalence classes of optimal lattice transformations, the 72 optimal changes of lattice basis split into three sets of 24 μ_{PQ}'s such that the 24 corresponding H_{PQ}'s lie in the same equivalence class.

4.2 Stability of the Bain strain

Theorem 2 showed that the Bain strain is optimal in an fcc-to-bcc phase transformation. In this section, we restrict our attention to $r = 1, 2$ and show that the Bain strain remains optimal for a range of lattice parameters in an fcc-to-bct phase transformation. This type of transformation is found in steels with higher carbon content. The strategy of the proof is to treat the bct phase as a perturbation of the bcc phase. To this end, for B as in (10), let the bct lattice be generated by

$$B_{\mathrm{AC}} = \mathrm{diag}(A, A, C)B = 2^{-4/3}\begin{pmatrix} -A & A & A \\ A & -A & A \\ C & C & -C \end{pmatrix}, \tag{16}$$

so that C denotes the elongation (or shortening) of the bcc cell in the z-direction and A the elongation (or shortening) in the x- and y-direction. We note that, since $\mathcal{P}^{24} = \mathrm{SL}(3,\mathbb{Z}) \cap \mathrm{SO}(3)$, the lattice $\mathcal{L}(B_{\mathrm{AC}})$ is equivalent to the lattices $\mathcal{L}(\mathrm{diag}(A, C, A)B)$ and $\mathcal{L}(\mathrm{diag}(C, A, A)B)$. Further we define $m_i^{0,\mathrm{AC}} = d_i(\mathrm{diag}(2^{1/6}A, 2^{1/6}A, 2^{-1/3}C), \mathbb{I})$.

The following proposition provides the most important ingredient.

Proposition 1. *("The first excited state")*
In a volume preserving transformation from an fcc-to-bcc lattice the second to minimal deformation distances are given by

$$m_1^1 := \min_{\substack{\mu \in \mathrm{SL}(3,\mathbb{Z}) \\ \mu \neq \mu_{\min}}} d_1(H_\mu, \mathbb{I}) \approx 0.70 \quad \textit{and} \quad m_2^1 := \min_{\substack{\mu \in \mathrm{SL}(3,\mathbb{Z}) \\ \mu \neq \mu_{\min}}} d_2(H_\mu, \mathbb{I}) \approx 1.64.$$

In particular, all H_μ with $\mu \in \mathrm{SL}(3,\mathbb{Z}) \backslash \mathrm{SL}^2(3,\mathbb{Z})$ have distance strictly larger than m_r^1, $r = 1, 2$.

Proof. For brevity let us call any deformation H_μ and the corresponding change of basis μ that has deformation distance m_r^1, $r = 1, 2$ an *excited state*. The proof follows along the same lines as the proof of Theorem 2. First we show with the help

of a computer that the second to minimal deformation distance within $\mathrm{SL}^2(3,\mathbb{Z})$ is given by the above and by (13) and (14) respectively applied on $\mathrm{SL}(3,\mathbb{Z})\backslash\mathrm{SL}^2(3,\mathbb{Z})$ we know that there cannot be any excited states in $\mathrm{SL}(3,\mathbb{Z})\backslash\mathrm{SL}^2(3,\mathbb{Z})$. □

Corollary 2. *("The first excited state" with volume change)*
In a transformation from an fcc-to-bcc lattice with volume change λ^3 the second to minimal deformation distances are given by

$$m_1^{1,\lambda} := \min_{\substack{\mu\in\mathrm{SL}(3,\mathbb{Z})\\ \mu\neq\mu_{\min}}} d_1(H_\mu^\lambda,\mathbb{I}) = 2^{-3/2}\sqrt{25\ 2^{1/3}\lambda^2 - 4\ 2^{2/3}\left(4+\sqrt{17}\right)\lambda + 24},$$
$$m_2^{1,\lambda} := \min_{\substack{\mu\in\mathrm{SL}(3,\mathbb{Z})\\ \mu\neq\mu_{\min}}} d_2(H_\mu^\lambda,\mathbb{I}) = 2^{-3}\ \sqrt{305\ 2^{2/3}\lambda^4 - 400\ 2^{1/3}\lambda^2 + 192}.$$

In particular all H_μ^λ with $\mu \in \mathrm{SL}(3,\mathbb{Z})\backslash\mathrm{SL}^2(3,\mathbb{Z})$ have distance strictly larger than $m_r^{1,\lambda}$, $r = 1, 2$.

Theorem 3. *The Bain strain is a d_1- and d_2-optimal lattice transformation from fcc-to-bct with lattice parameters A, C in the range*

$$\{(A,C) : C \geq A > 0.75 \text{ and } m_1^1 - 3^{3/2}|B_{AC} - B| \geq m_1^{0,AC}\} \qquad \text{for } r = 1,$$
$$\{(A,C) : C \geq A > 0.75 \text{ and } m_2^1 - 27|B_{AC}^{\mathrm{T}}B_{AC} - B^{\mathrm{T}}B| \geq m_2^{0,AC}\} \qquad \text{for } r = 2,$$

(cf. Figure 3). For $C > A$ the stretch components of the optimal lattice transformations are given by $\bar{H}_{\min}^{AC}$ in the set

$$\left\{\begin{pmatrix} 2^{1/6}A & 0 & 0\\ 0 & 2^{1/6}A & 0\\ 0 & 0 & 2^{-1/3}C\end{pmatrix}, \begin{pmatrix} 2^{1/6}A & 0 & 0\\ 0 & 2^{-1/3}C & 0\\ 0 & 0 & 2^{1/6}A\end{pmatrix}, \begin{pmatrix} 2^{-1/3}C & 0 & 0\\ 0 & 2^{1/6}A & 0\\ 0 & 0 & 2^{1/6}A\end{pmatrix}\right\}.$$

The respective minimal metric distances are

$$m_{\min,1}^{AC} = d_1(H_{\min}^{AC},\mathbb{I}) = \left(2\left(2^{1/6}A - 1\right)^2 + \left(2^{-1/3}C - 1\right)^2\right)^{1/2} = m_1^{0,AC}, \tag{17}$$
$$m_{\min,2}^{AC} = d_2(H_{\min}^{AC},\mathbb{I}) = \left(2\left(\left(2^{1/6}A\right)^2 - 1\right)^2 + \left(\left(2^{-1/3}C\right)^2 - 1\right)^2\right)^{1/2} = m_2^{0,AC} \tag{18}$$

and are achieved by exactly 24 distinct $\mu \in \mathrm{SL}(3,\mathbb{Z})$. The case $C = A$ corresponds to an fcc-to-bcc transformation with volume change λ^3 with $\lambda = A = C$ and we refer to Corollary 1.

Example 4. $A = C = 1$ recovers Theorem 2. $C > A$ corresponds to the usual fcc-to-bct transformation for steels with higher carbon content. $C = \sqrt{2}A = 2^{1/3}$ is the bct lattice that is contained in the fcc lattice, i.e. $d(\mathcal{L}_0, \mathcal{L}_1) = 0$.

Proof (of Theorem 3). We will show that precisely 24 of the 72 μ's that were optimal in the fcc-to-bcc transition remain optimal. Let us start from one of those optimal

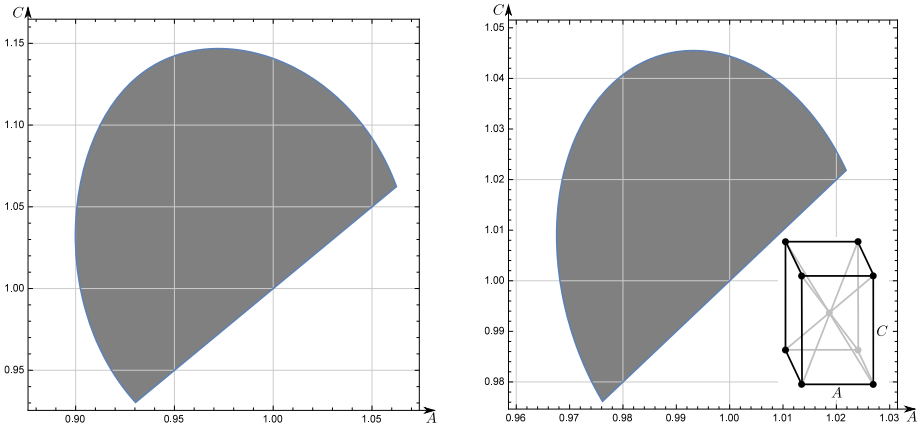

Figure 3: The range of A, C where the Bain strain remains d_1-optimal (left) and d_2-optimal (right).

transformations from fcc-to-bcc given by e.g.

$$\mu_0 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

We know that $H^{\mathrm{AC}}_{\mu_0} = B_{\mathrm{AC}}\mu_0 F^{-1}$ is optimal for $A = C = 1$, where B_{AC} is given by (16). The deformation distance is $d_r(H^{\mathrm{AC}}_{\mu_0}, \mathbb{I}) = m^{\mathrm{AC}}_{\mathrm{min},r}$ with $m^{\mathrm{AC}}_{\mathrm{min},r}$ given by (17) and (18) respectively. Further there exist 24 different μ's and corresponding H^{AC}_{μ}'s that have the same distance. This follows as in Remark 7 with the only difference that the point group of the bct lattice only has 8 elements. Then, due to the invariance of d_1 and d_2 under multiplication from the left or right by any rotation, the 24 matrices H^{AC}_{μ} trivially have the same distance to $\mathbb{I}$ and one may easily verify that these H^{AC}_{μ}'s are equally split into the three equivalence classes as in the statement of the Theorem.

The remaining 48 μ's that were optimal for fcc-to-bcc lead to a larger deformation distance. One of these non-optimal lattices is generated by

$$\tilde{B}_{\mathrm{AC}} = 2^{-4/3} \begin{pmatrix} -A & A & 0 \\ A & A & 0 \\ -C & -C & -2C \end{pmatrix}$$

with

$$\left(d_1(\tilde{B}_{\mathrm{AC}}F^{-1}, \mathbb{I})\right)^2 - \left(m^{\mathrm{AC}}_{\mathrm{min},1}\right)^2 = 2^{-5/3}\,(C - A)\left(2^{1/6}\,A^2 + 2^{1/6}\,C^2 - 4 + 2^{3/2}\right) > 0$$

$$\Leftarrow\ C > A \text{ and } A^2 + C^2 > \frac{2^{5/6}}{2 - \sqrt{2}} \quad \Leftarrow\ C > A > 0.75 \tag{19}$$

and

$$\left(d_2(\tilde{B}_{\mathrm{AC}}F^{-1},\mathbb{I})\right)^2-\left(m^{\mathrm{AC}}_{\mathrm{min},2}\right)^2=2^{-4/3}\,(C-A)\,(A+C)\left(3\,A^2+3\,C^2-2\,2^{2/3}\right)>0$$
$$\Leftarrow\ C>A \text{ and } A^2+C^2>\frac{2^{8/3}}{3}\quad\Leftarrow\ C>A>0.75 \tag{20}$$

which holds true for all $C>A$ in the range under consideration. The remaining 47 non-optimal deformations H^{AC}_μ are all $\mathcal{P}^{24}$ related and thus have the same distance; in particular larger than $m^{\mathrm{AC}}_{\mathrm{min},r}$.

To show the minimality of the 24 H^{AC}_μ's, we make use of our result on the first excited state to compare their distance to $\mathbb{I}$ against all the remaining μ's that were non-optimal in the fcc-to-bcc transition. In particular, we need to show that

$$m^{\mathrm{AC}}_{\mathrm{min},i}<\min_{\mu\neq\mu_{\mathrm{min}}} d_r(H^{\mathrm{AC}}_\mu,\mathbb{I}),\quad r=1,2,$$

where $H^{\mathrm{AC}}_\mu=B_{\mathrm{AC}}\mu F^{-1}$ and μ_{min} is any of the 72 minimising μ's from Theorem 2. Let us set $H_\mu:=H^{11}_\mu$ and estimate using the properties of d_r (cf. Example 1)

$$\begin{aligned}\min_{\mu\neq\mu_{\mathrm{min}}} d_r(H^{\mathrm{AC}}_\mu,\mathbb{I})&\geq\min_{\mu\neq\mu_{\mathrm{min}}}\left(d_r(H_\mu,\mathbb{I})-d_r(H^{\mathrm{AC}}_\mu,H_\mu)\right)\\&\geq m^1_r-\max_{\mu\neq\mu_{\mathrm{min}}} d_r(H^{\mathrm{AC}}_\mu,H_\mu),\end{aligned}\tag{21}$$

where m^1_r denotes the first excited state (cf. Proposition 1). We estimate

$$\begin{aligned}&d_1(H^{\mathrm{AC}}_\mu,H_\mu)\leq|H^{\mathrm{AC}}_\mu-H_\mu|\leq|B_{\mathrm{AC}}-B||\mu F^{-1}|,\\&d_2(H^{\mathrm{AC}}_\mu,H_\mu)\leq|H^{\mathrm{AC}\mathrm{T}}_\mu H^{\mathrm{AC}}_\mu-H^{\mathrm{T}}_\mu H_\mu|\leq|B^{\mathrm{T}}{}_{\mathrm{AC}}B_{\mathrm{AC}}-B^{\mathrm{T}}B||\mu F^{-1}|^2\end{aligned}$$

and with the help of a computer we calculate $\max_{\mu\in\mathrm{SL}^1(3,\mathbb{Z})}|\mu F^{-1}|=3^{3/2}$. Thus a sufficient condition that the Bain strain is d_1-optimal within $\mathrm{SL}^1(3,\mathbb{Z})$ is that $0.75<A\leq C$ satisfy

$$m^1_1-3^{3/2}|B_{\mathrm{AC}}-B|\geq m^{\mathrm{AC}}_{\mathrm{min},1}$$

yielding the area drawn in Figure 3 (left). To exclude any $\mu\in\mathrm{SL}(3,\mathbb{Z})\backslash\mathrm{SL}^1(3,\mathbb{Z})$ we replace b by $b_{\mathrm{AC}}=\mathrm{diag}(A,A,C)b$ in (12) in the proof of Theorem 2 and estimate $|b_{\mathrm{AC}}|\geq A|b|$. Concluding as in (13) we arrive at

$$\min_{\mu\in\mathrm{SL}(3,\mathbb{Z})\backslash\mathrm{SL}^1(3,\mathbb{Z})} d_1(H^{\mathrm{AC}}_\mu,\mathbb{I})\geq 2^{2/3}A-1,\tag{22}$$

which needs to be larger than $m^{\mathrm{AC}}_{\mathrm{min},1}$. This holds true e.g. for $A\in[0.85,1.7]$ and $C\in[A,1.7]$. The proof of the d_2-optimality proceeds analogously. To obtain d_2-optimality within $\mathrm{SL}^1(3,\mathbb{Z})$ we need to satisfy

$$m^1_2-27|B^{\mathrm{T}}_{\mathrm{AC}}B_{\mathrm{AC}}-B^{\mathrm{T}}B|\geq m^{\mathrm{AC}}_{\mathrm{min},2}$$

for all $0.75 < A \leq C$ which yields the area drawn in Figure 3 (right). To exclude all elements in the complement of $\mathrm{SL}^1(3,\mathbb{Z})$ we have to ensure that $2^{4/3}A - 1 > m^{\mathrm{AC}}_{\mathrm{min},2}$ which holds true e.g. for $A \in [0.75, 1.5]$ and $C \in [A, 1.5]$. □

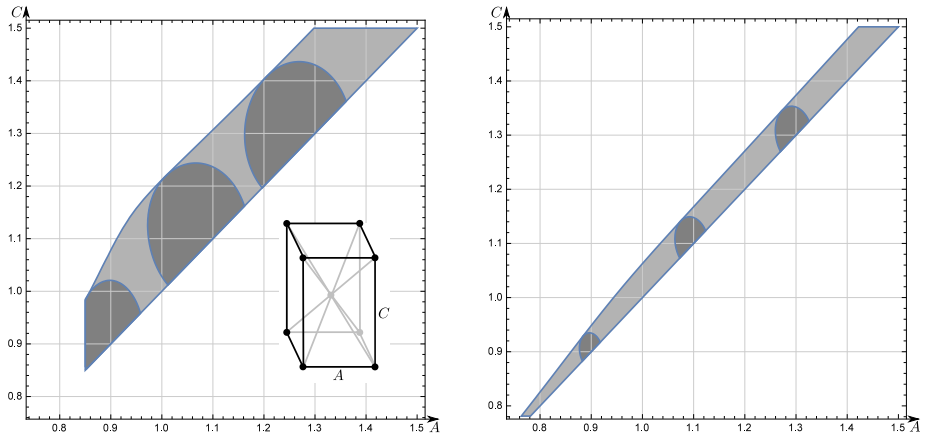

Figure 4: Extended d_1-optimality (left) and d_2-optimality (right) range for A and C. The dark shaded regions are obtained by a perturbation argument about the fixed optimal transformations $H^{\lambda_0}_{\mathrm{min}}$, $\lambda_0 = 0.9, 1.1, 1.3$. The light shaded regions are obtained by a perturbation argument about the (A,C)-dependent optimal transformation $H^{\lambda(A,C)}_{\mathrm{min}}$ with $\lambda(A,C) = 0.995\sqrt{AC}$.

Remark 8. By employing Corollary 1 and Corollary 2 we may extend the range of optimal parameters A and C in Theorem 3 (cf. Figure 3) by shifting the reference point from $A = C = 1$ to $A = C = \lambda$. This enables us to show that the Bain strain remains the d_1- and d_2-optimal lattice deformation from fcc-to-bct in a much larger range of lattice parameters $C \geq A$. The shaded regions in Figure 4 show the values of A, C such that the deformation $H^{\mathrm{AC}}_{\mathrm{min}}$ remains d_1- and d_2-optimal.

Proof. The main idea of the proof is to consider the element $H^{\lambda}_{\mu} = \lambda H_{\mu}$ which is optimal in an fcc-to-bcc transformation with volume change (cf. Corollary 1) and "closest" to H^{AC}_{μ}. Thus in (21) we write

$$\min_{\mu \neq \mu_{\mathrm{min}}} d_r(H^{\mathrm{AC}}_{\mu}, \mathbb{I}) \geq \min_{\mu \neq \mu_{\mathrm{min}}} \left(d_r(H^{\lambda}_{\mu}, \mathbb{I}) - d_r(H^{\lambda}_{\mu}, H^{\mathrm{AC}}_{\mu}) \right)$$

and for d_1 we estimate

$$\min_{\mu \neq \mu_{\mathrm{min}}} d_1(H^{\mathrm{AC}}_{\mu}, \mathbb{I}) \geq m^{1,\lambda}_1 - \max_{\mu \neq \mu_{\mathrm{min}}} \left| \lambda H_{\mu} - H^{\mathrm{AC}}_{\mu} \right| \geq m^{1,\lambda}_1 - \lambda \max_{\mu \neq \mu_{\mathrm{min}}} |\mu F^{-1}| \|B - B_{\frac{A}{\lambda}\frac{C}{\lambda}}\|.$$

To obtain the d_1-optimality in $\mathrm{SL}^1(3,\mathbb{Z})$, we again use $\max_{\mu \in \mathrm{SL}^1(3,\mathbb{Z})} |\mu F^{-1}| = 3^{3/2}$ so that $m^{1,\lambda}_1 - \lambda 3^{3/2} \|B - B_{\frac{A}{\lambda}\frac{C}{\lambda}}\| \geq m^{\mathrm{AC}}_{\mathrm{min},1}$. Condition (22) regarding d_1-optimality outside $\mathrm{SL}^1(3,\mathbb{Z})$ remains unchanged. Both conditions combined yield the area

shown in Figure 4 (left). Analogously in order to show d_2-optimality we estimate

$$\begin{aligned}\min_{\mu\neq\mu_{\min}} d_2(H_\mu^{\mathrm{AC}},\mathbb{I}) &\geq m_2^{1,\lambda} - \max_{\mu\neq\mu_{\min}} \left|\lambda^2 H_\mu^{\mathrm{T}} H_\mu - H_\mu^{\mathrm{AC}\,\mathrm{T}} H_\mu^{\mathrm{AC}}\right| \\ &\geq m_2^{1,\lambda} - \lambda^2 \max_{\mu\neq\mu_{\min}} |\mu F^{-1}|^2 \left|B^{\mathrm{T}}B - B^{\mathrm{T}}_{\frac{A}{\lambda}\frac{C}{\lambda}} B_{\frac{A}{\lambda}\frac{C}{\lambda}}\right|\end{aligned}$$

and again use $\max_{\mu\in\mathrm{SL}^1(3,\mathbb{Z})} |\mu F^{-1}| = 3^{3/2}$ to show d_2-optimality within $\mathrm{SL}^1(3,\mathbb{Z})$. The condition to be d_2-optimal outside $\mathrm{SL}^1(3,\mathbb{Z})$ remains unchanged. Both conditions combined yield the area shown in Figure 4 (right). □

Remark 9. The previous estimates can be iterated, i.e. instead of picking the H_μ^λ that is closest to H_μ^{AC} one may pick any $H_\mu^{\mathrm{A'C'}}$ that is in the range indicated in Figure 4 that is closest to H_μ^{AC}.

If $C \leq A$, following the proof of Theorem 3, one finds that the optimal strain becomes

$$\bar{H}_{\min} = \begin{pmatrix} 2^{1/6}\frac{A+C}{2} & \pm 2^{1/6}\frac{A-C}{2} & 0 \\ \pm 2^{1/6}\frac{A-C}{2} & 2^{1/6}\frac{A+C}{2} & 0 \\ 0 & 0 & 2^{-1/3}A \end{pmatrix} + \text{ its } \mathcal{P}^{24} \text{ conjugates}$$

at least if A and C are in the regions specified in the statement of Theorem 3 with $C \geq A$ replaced by $A \geq C$.

4.3 Terephthalic Acid

Terephthalic Acid is a material that has two triclinic phases (Type I and II) which are very different from each other (cf. [BJ15, p.46 ff.]). Thus any lattice transformation necessarily requires large principal stretches and, unlike in the Bain setting, it is not clear what a good candidate for the optimal transformation would be. However, with the help of the proposed framework the d_r-optimal lattice transformation can easily be determined. The only required input parameters are the lattice parameters of the two triclinic unit cells (cf. [BB67, p.388 Table 2]) listed in Table 3.

Form	$a/A°$	$b/A°$	$c/A°$	$\alpha/°$	$\beta/°$	$\gamma/°$
I	7.730	6.443	3.749	92.75	109.15	95.95
II	7.452	6.856	5.020	116.6	119.2	96.5

Table 3: Lattice parameters of the triclinic unit cells of Terephthalic Acid

To apply our analysis we first ought to convert the triclinic to the primitive description.

Lemma 7. *(Conversion from triclinic to primitive unit cell)*
The triclinic unit cell with lattice parameters a, b, c and α, β, γ generates up to an

overall rotation the same lattice as

$$F = \begin{pmatrix} a & b\cos(\gamma) & c\cos(\beta) \\ 0 & b\sin(\gamma) & c\sin^{-1}\gamma\left(\cos(\alpha)-\cos(\beta)\cos(\gamma)\right) \\ 0 & 0 & c\left(\sin^2\beta-\sin^{-2}\gamma\left(\cos(\alpha)-\cos(\beta)\cos(\gamma)\right)^2\right)^{1/2} \end{pmatrix}$$

Proof. Let $\mathrm{col}[F] = \{f_1, f_2, f_3\}$. It is easy to verify that $|f_1| = a$, $|f_2| = b$, $|f_3| = c$ and that $\measuredangle(f_1, f_2) = \gamma$, $\measuredangle(f_1, f_3) = \beta$, $\measuredangle(f_2, f_3) = \alpha$. □

Example 5. (Primitive cells of Terephthalic Acid)
Application of the previous Lemma to the lattice parameters in Table 3 leads to

$$F_I = \begin{pmatrix} 7.730 & -0.668 & -1.230 \\ 0 & 6.408 & -0.309 \\ 0 & 0 & 3.528 \end{pmatrix} \text{ and } F_{II} = \begin{pmatrix} 7.452 & -0.776 & -2.449 \\ 0 & 6.812 & -2.541 \\ 0 & 0 & 3.570 \end{pmatrix} \quad (23)$$

(all measures in A°)

Theorem 4. *(Optimal lattice transformations in Terephthalic Acid)*
The unique equivalence class of d_1- and d_2-optimal transformations between Terephthalic Acid Form I and Terephthalic Acid Form II has a stretch component given by

$$\bar{H}_{\min} = \begin{pmatrix} 0.820 & -0.125 & -0.072 \\ -0.125 & 0.994 & -0.146 \\ -0.072 & -0.146 & 1.329 \end{pmatrix} \quad (24)$$

with principal stretches $\nu_1 = 0.725$, $\nu_2 = 1.033$ and $\nu_3 = 1.385$. The minimal distances are given by $m_{\min,1} = d_1(\bar{H}_{\min}, \mathbb{I}) = 0.474$ and $m_{\min,2} = d_2(\bar{H}_{\min}, \mathbb{I}) = 1.035$.

Proof. We apply Theorem 1 with $F = F_I$ and $G = F_{II}$. We calculate $\|F_I\|_{2,\infty} = |F_I e_1| = 7.730 < 7.8$, $\nu_{\min}(F_{II}) = 3.076 > 3$, $m_{0,1} = 0.529 < 0.55$ and $m_{0,2} = 1.197 < 1.2$. Further we note that if $\mu \in \mathrm{SL}(3,\mathbb{Z})\backslash\mathrm{SL}^{k-1}(3,\mathbb{Z})$ then $\|\mu\|_{2,\infty} \geq \sqrt{k^2+1}$. Therefore by (8) any d_1-optimal μ satisfies

$$\|\mu\|_{2,\infty} \leq \frac{\|F\|_{2,\infty}}{\nu_{\min}(G)}(m_{0,1}+1) < \frac{7.8}{3}(0.55+1) = 4.03 < \sqrt{4^2+1}$$

and is thus contained in $\mathrm{SL}^3(3,\mathbb{Z})$ and by (8) any d_2-optimal μ satisfies

$$\|\mu\|_{2,\infty}^2 \leq \frac{\|F\|_{2,\infty}^2}{\nu_{\min}^2(G)}(m_{0,2}+1) < \frac{7.8^2}{3^2}(1.2+1) = 14.872$$

and is thus also contained in $\mathrm{SL}^3(3,\mathbb{Z})$. With the help of a computer we check that within the set $\mathrm{SL}^3(3,\mathbb{Z})$ the minimum is in both cases attained at

$$\mu_{\min} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & -1 \end{pmatrix}$$

with corresponding $\bar{H}_{\min} = \sqrt{H^{\mathrm{T}}{}_{\mu_{\min}} H_{\mu_{\min}}}$ as in (24), $m_{\min,1} = d_1(\bar{H}_{\min}, \mathbb{I}) = 0.474$ and $m_{\min,2} = d_2(\bar{H}_{\min}, \mathbb{I}) = 1.035$. □

Remark 10. We have shown that the d_1- and d_2-optimal transformations from Form I to Form II of Terephthalic Acid are the same. However the d_{-2}-optimal transformation is different and given by

$$\bar{H}_{\min,-2} = \begin{pmatrix} 0.852 & -0.119 & -0.018 \\ -0.119 & 0.950 & -0.197 \\ -0.018 & -0.197 & 1.346 \end{pmatrix}$$

with principal stretches $\nu_1 = 0.743$, $\nu_2 = 0.977$ and $\nu_3 = 1.429$. As expected the smallest principal stretch ν_1 is bigger than before since d_{-2} penalises contractions significantly more than expansions. To obtain the required analytical bounds one calculates $\|F_{II}\|_{2,\infty} = |F_{II} e_1| = 7.452 < 7.46$, $\nu_{\min}(F_I) = 3.464 > 3.45$ and $m_{0,-2} = 1.080 < 1.1$ to get $\|\mu^{-1}\|^2_{2,\infty} < 10$ (cf. (9)) and thus the optimal μ lies in $\mathrm{SL}^{-2}(3,\mathbb{Z})$.

Remark 11. Even though [CSTJ16] also uses the distance d_{-2}, the reported strains are different. However, this discrepancy is not a consequence of using sublattices but rather a consequence of using the lattice parameters from [BB67, p.388 Table 1] for Terephthalic Acid II[3] and the lattice parameters from [BB67, p.388 Table 2] for Terephthalic Acid I. In particular, using these mixed input parameters in "OptLat" (cf. Appendix) yields identical values.

5 Concluding remarks

The present paper provides a rigorous proof for the existence of an optimal lattice transformation between any two given Bravais lattices with respect to a large number of optimality criteria. Furthermore, a precise algorithm and a GUI to determine this optimal transformation is provided (cf. Appendix). As possible applications, the optimal transformation in steels, that is the transformation from fcc-to-bcc/bct, and in Terephthalic Acid were determined. Through Theorem 1 and with the help of the provided algorithm/programme one is able to rigorously determine the optimal phase transformation in any material undergoing a displacive phase transformation from one Bravais lattice to another.

If the parent or product phases are multilattices, the proposed framework is not *a priori* applicable. Nevertheless, one may still consider Bravais sublattices of these multilattices and proceed as before. The choice of these sublattices may come from physical consideration. However, in order to rigorously determine the optimal transformation between two given multilattices one would need to measure the movement of *all* atoms consistently, that is one would need to take into account both the overall periodic deformation of the unit cell and the shuffle movement of atoms within the unit cell. Establishing such a criterion would be of great interest but lies outside the scope of the present paper.

[3] $a = 9.54A°$, $b = 5.34A°$, $c = 5.02A°$, $\alpha = 86.95°, \beta = 134.65°, \gamma = 94.8°$

Appendix

A. Mathematica

The following Mathematica code[4] determines for a given $k \in \mathbb{N}$ the optimal lattice transformation $H : \mathcal{L}(F) \to \mathcal{L}(G)$ within the set $\{H_\mu = G\mu F^{-1} : \mu \in \mathrm{SL}^k(3,\mathbb{Z})\}$ for any given $F, G \in \mathrm{GL}^+(3,\mathbb{R})$ and for any distance measure $d_r(H,\mathbb{I})$, $r > 0$. The case $r < 0$ is analogous.

For the transformation from fcc-to-bcc, F and G would be given by (10) and for the transformation from Terephthalic Acid I to II, F and G would be given by (23).

Firstly, we generate the set $\mathrm{SL}^k(3,\mathbb{Z})$ (=SL):

```
SL = Select[Flatten[Table[{a,b,c,d,e,f,g,h,i},
      {a,-k,k},{b,-k,k},{c,-k,k},{d,-k,k},{e,-k,k},{f,-k,k},
      {g,-k,k},{h,-k,k},{i,-k,k}],8],Det[Partition[#,3]]==1&];
```

Next we generate a list (=distlist) of all values of $d_r(H_\mu,\mathbb{I})$ for $\mu \in \mathrm{SL}^k(3,\mathbb{Z})$:

```
Hmu = Function[mu,G.Partition[mu,3].Inverse[F]];
distr = Function[mu,Norm[SingularValueList[Hmu[mu]]^r-{1,1,1}]];
distlist = distr/@SL;
```

Then we generate a list (=poslist) of all the positions of μ's in $\mathrm{SL}^k(3,\mathbb{Z})$ that give rise to the minimal deformation distance:

```
poslist = Flatten[Position[distlist,RankedMin[distlist,1]],1];
```

Further we calculate the minimal deformation distance m_0, the second to minimal deformation distance m_1 and return their numerical difference $\Delta = m_1 - m_0$ (=delta):

```
m0 = distlist[[poslist[[1]]]]; m1 = Sort[distlist][[Length[poslist]+1]];
delta = N[m1-m0]
```

Finally we return a list of all μ's that give rise to an optimal deformation H_μ and a list of all optimal H_μ's:

```
SL[[poslist]]
Hmu/@SL[[poslist]]
```

B. MATLAB

A Graphical User Interface (GUI) called "OptLat" can either be found on MATLAB File Exchange[5] (requires MATLAB) or downloaded directly as a standalone Windows application[6].

Acknowledgements: We would like to thank R. D. James and X. Chen for helpful discussions. The research of A. M. leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 291053.

[4] The original .nb file can be found online under http://solids.maths.ox.ac.uk/programs/OptLat.nb
[5] http://uk.mathworks.com/matlabcentral/fileexchange/55554-optlat
[6] http://solids.maths.ox.ac.uk/programs/OptLat.exe

References

[Bai24] E.C. Bain. The nature of martensite. *Transactions of the Metallurgical Society of AIME*, 70(1):25–47, 1924.

[BB67] M. Bailey and C.J. Brown. The crystal structure of Terephthalic Acid. *Acta Crystallographica*, 22(3):387–391, 1967.

[Bha01] H.K.D.H. Bhadeshia. *Geometry of crystals*. Institute of Materials, London, 2001.

[Bha03] K. Bhattacharya. *Microstructure of martensite*. OUP Oxford, 2003.

[BJ92] J.M. Ball and R.D. James. Proposed experimental tests of a theory of fine microstructure and the two-well problem. *Philosophical Transactions: Physical Sciences and Engineering*, pages 389–450, 1992.

[BJ15] J.M. Ball and R.D. James. Incompatible sets of gradients and metastability. *Archive for Rational Mechanics and Analysis*, 218(3):1363–1416, 2015.

[BW72] J.S. Bowles and C.M. Wayman. The Bain strain, lattice correspondences, and deformations related to martensitic transformations. *Metallurgical Transactions*, 3(5):1113–1121, 1972.

[CSTJ16] Xian Chen, Yintao Song, Nobumichi Tamura, and Richard D. James. Determination of the stretch tensor for structural transformations. *Journal of the Mechanics and Physics of Solids*, pages –, 2016.

[DE56] T.C. Doyle and J.L. Eriksen. Non-linear elasticity. In *Advances in Applied Mechanics*, volume 4, pages 53–115. Elsevier, 1956.

[Gen07] J.E. Gentle. *Matrix Algebra: Theory, Computations, and Applications in Statistics*. Springer Texts in Statistics. Springer New York, 2007.

[JW48] M.A. Jaswon and J.A. Wheeler. Atomic displacements in the austenite–martensite transformation. *Acta Crystallographica*, 1(4):216–224, 1948.

[KM] K. Koumatos and A. Muehlemann. A theoretical investigation of orientation relationships and transformation strains in steels.

[KS30] G.V. Kurdjumov and G. Sachs. Über den Mechanismus der Stahlhärtung. *Zeitschrift für Physik*, 64:325–343, 1930.

[Lom55] W.M. Lomer. The $\beta \to \alpha$ transformation in Uranium-1.4 at.% Chromium alloy. *Inst. Metals, London Monograph No. 18*, pages 243–253, 1955.

[PZ02] M. Pitteri and G. Zanzotto. *Continuum models for phase transitions and twinning in crystals*. CRC Press, 2002.